

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

## ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Cohort Profile: Follow-up of a Berlin Aging Study II (BASE-II) subsample as part of the GendAge study
<b>AUTHORS</b>	Demuth, Ilja; Banszerus, Verena; Drewelies, Johanna; Düzel, Sandra; Seeland, Ute; Spira, Dominik; Tse, Esther; Braun, Julian; Steinhagen-Thiessen, Elisabeth; Bertram, Lars; Thiel, Andreas; Lindenberger, Ulman; Regitz-Zagrosek, Vera; Gerstorf, Denis

## VERSION 1 – REVIEW

<b>REVIEWER</b>	Ferrannini, Giulia Karolinska Institute
<b>REVIEW RETURNED</b>	03-Jan-2021

<b>GENERAL COMMENTS</b>	<p>The authors performed the GendAge study, focusing on sex and gender differences, in a cohort of patients from the BASE-II study.</p> <p>This study had 2 phases:</p> <ul style="list-style-type: none"><li>- a recruitment phase (2009-2014)</li><li>- a follow-up phase (2018-2020)</li></ul> <p>including two different age groups, if my understanding is correct:</p> <ul style="list-style-type: none"><li>- an older age group (1671 at baseline and 1100 at follow up with complete medical information)</li><li>- a younger age group (500 or 600 at baseline - not clear - and 64 at follow up)</li></ul> <p>I found this population selection not easy to understand. Moreover, it was not possible to complete the follow up of the younger participants due to COVID-19, and accordingly, there are no results described for them nor there is a table, but some scattered information is given in the text.</p> <p>I wonder whether it wouldn't be better to describe just the older population or, alternatively, to wait to have the younger population assessments completed.</p> <p>The methods are well described and accurate, specifically as regards the gender-centered questionnaire, and there is a lot of useful information provided, which adds to the value of this study.</p> <p>However, I think that the difference between gender-related variables and sex-related assessments should be more clearly distinguished, e.g. in two different paragraphs; this would help the reader to distinguish between the two sets of variables and encourage their proper classification. There are factors, indeed, which are hard to uniquely attribute to gender or sex, but still, this should be discussed.</p>
-------------------------	--

	<p>Another issue is that, if I understood correctly, the authors are describing the older group only (Table 1), as in the younger group follow-up was not completed. This causes some confusion in the manuscript and does not really help to focus on the major issue.</p> <p>The results should be more thoroughly described and, even though this is just a cohort profiling, might start to be put in relation to previous findings on the matter.</p> <p>A major limitation is that the numbers are quite low. Surely, to make a great number of detailed assessments in very large populations is quite challenging, but this should nevertheless be highlighted in the limitations.</p>
--	--

<b>REVIEWER</b>	Holmberg, Mats Karolinska Universitetssjukhuset, ANOVA
<b>REVIEW RETURNED</b>	14-Feb-2021

<b>GENERAL COMMENTS</b>	<p>To begin: I am not familiar with this type of manuscript, in fact, this is my first encounter with a manuscript of this type. But as I understand it, it should be a thorough description of a cohort and what the plans and considerations are for this cohort. This is what I have had in mind while reviewing.</p> <p>I also want to stress that I am not a native English speaker. Both these circumstances may colour my views.</p> <p>Having said this my first suggestion will be to have this manuscript language checked. I believe many readers will be unfamiliar with the structure of the manuscript and correct English will enhance understanding.</p> <p>The second will be: stay with your object, describe the collection of the cohort, how was it done, with what purpose, how will it be investigated and what are the advantages and disadvantages of the choices you've made? I understand that you already have started to perform the analysis you planned to do but reporting results in this type of manuscript is, unfortunately, confusing. I will not go into much detail now since I feel that you need to make a choice on what type of manuscript you want to write and stick to that.</p> <p>When commenting I use the paging on the top of each page.</p> <p>Specific comments</p> <p>Title: By the title, it seems to me that you are presenting two cohorts? Or a study and a cohort. Or should the title be: Cohort profile: the "older BASE II"</p> <p>I will come back to this issue many times. Which cohort do you want to present and my guess is that you want to present the creation of the "older BASE II", a subsample of the BASE II study. Or the "GendAge study cohort"? the follow-up of a subsample from the BASE II study.</p> <p>Abstract</p> <p>Purpose: Are you presenting a study (a protocol) or a cohort?. To me this is confusing. It seems that you start by presenting a study.</p> <p>Participants: Is this the cohort? A total of 1,100 participants (older BASE-II group) This is coherent with Table 1.</p>
-------------------------	---

	<p>Strengths and limitations I think this section should be about the strengths and limitations of how you have designed your cohort, it includes too little of this and too much of other information now.</p> <p>Introduction What is the reason for the new data collection? P6r58: you are here referring to the Base II cohort but at the same time it seems as you are referring to the "older BASE II". If you want to report on different cohorts I think it is very important that you use different names consistently. P7r3: This new instrument seems interesting and I would appreciate a more thorough description of the instrument and its advantages and disadvantages P8r13 The numbers "the remaining pool of 1,428 subjects out of the originally 1,671" is a little confusing since they don't match the previously described numbers. P8r28: So the cohort was not only formed from the BASE II? You added 17 more? See below. P8r40: the younger ones? I thought this cohort was the older ones? Confusing P9r6: When was the follow-up? How long time? You mention how long time it took but what was the plan? Findings to date I am sorry, but this section does seem like a study report, and as I understand this type of publication it should describe the cohort, how it was collected and what you plan to do with it. So I simply don't think this section belongs in this article. Table 1 – I find this table clear in the description of the characteristics of the cohort. Under p-value, I think it's better just to use &lt;0.05 or &lt; 0.001 instead of publishing extremely low numbers. They are of no practical value. Table 2 – I don't think the plan for the three visits was described before? Figure 1 – I find it hard to understand the relationship between the 17 and the rest in the BASE II cohort. You mention 1671+500 who were medically assessed. That is 2171. 29 are missing up to 2200. Was the 17 part of these 29? What about the remaining 12?</p>
<b>REVIEWER</b>	Haseeb, Sohaib James Cook University, College of Medicine and Dentistry
<b>REVIEW RETURNED</b>	18-Feb-2021
<b>GENERAL COMMENTS</b>	<p>The GendAge study collects follow-up data from the BASE-II study focusing on cardiovascular and metabolic diseases in a broad cohort of volunteers from Berlin, Germany. An array of data are being collected, including baseline characteristics, echocardiography data, laboratory parameters, and psychosocial and cognitive data. A gender score has also been reportedly developed.</p> <p>Strengths of this paper include the broad and comprehensive nature of the data collection, large sample size constituting healthy volunteers, and comparison of age- gender- and sex-related differences. The introduction appropriately frames the study.</p>

	<p>Limitations of this paper include the broad range of outcomes being analyzed. It is unclear what the specific objectives are and which data have been collected and are presented as a finding. As such, it is tedious to weigh the entire piece together.</p> <p>Specific Comments:</p> <p>In general, the section "Findings to Date" does not present the data completely. Instead, it appears that more of a methodology is presented rather than findings to date. It is unclear which data are presented and which are still being collected. Specifically:</p> <p>Page 9, Lines 21-34: A description of the gender questionnaire and score are given. The score and/or the elements constituting it have not been presented. Neither has it been applied and validated in the current population. Therefore, I am confused as to how this applies to the current study.</p> <p>Can the elements of the score be presented and applied to the participants? A separate table can be prepared to present this data.</p> <p>Page 9, Lines 21-34: It is suggested that data on atrial and LV morphology, systolic and diastolic function, and vascular stiffness were obtained. Although, I can only see the data for LVEF in Table 1 and its appropriate comparison in the text; other values have not been compared.</p> <p>If possible, it would be beneficial to report other echocardiography parameters and their gender-related comparisons.</p> <p>Furthermore, I am curious why ICCs between an analyst and an independent supervisor are emphasized? This is an important quality-control mechanism, but perhaps not a finding.</p> <p>Page 10, Other measurements: Where is the data concerning the psychological questionnaire and the cognitive assessments? The authors have stated that these assessments were completed, but to my eyeball review, no data have been compared — at least in the text. Are they part of Table 1?</p> <p>Page 11, Lines 3-5: Similarly, it is stated that data concerning participant's physical activity and sleep were collected. Where are they presented? It is unclear whether these data were part of this study. For instance, much research has been performed on sleep apnea and atrial fibrillation. If such a relationship exists here would be interesting to know.</p> <p>Page 8-11: Is there a possibility for the authors to separate the methodology from the findings. This section is quite convoluted. It starts off by presenting relevant baseline characteristics but quickly resorts to integrating methodologic elements with the findings.</p>
--	--

	<p>Perhaps a separate section is needed for: (1) Gender questionnaire and score description w/ its elements; (2) Echocardiographic measurements; (3) Collection of genomic samples; (4) Blood sampling; and (5) Ethics clearance.</p> <p>I congratulate the authors on conducting a comprehensive longitudinal study of this magnitude involving participants from the community.</p>
--	---

<b>REVIEWER</b>	Heer, Tobias Klinikum Schwabing, Cardiology
<b>REVIEW RETURNED</b>	23-Feb-2021

<b>GENERAL COMMENTS</b>	<p>The authors describe the cohort profile of the GendAge study, which is based on a reinvestigation of the Berlin Aging Study II (BASE-II). The present study focuses on major risk factors for cardiovascular and metabolic diseases and on the development of major outcomes from intermediate phenotypes in the context of biological sex and gender differences.</p> <p>1. The authors should clarify if the GendAge cohort is a subgroup of the BASE-II study, which means that it is not a different population. Is the BASE-II study separately ongoing and is the GendAge population still integrated in the BASE-II study? Is GendAge the follow-up study of BASE-II? The difference of the two studies should be clarified. What does reinvestigation mean?</p> <p>2. In the abstract the authors write that the gender score was associated with a number of [...] variables and performed better in predicting differences in a subset of variables. The authors should clarify what differences they are talking about.</p> <p>3. Is there also a group with intervention to enhance healthy ageing or is this an observational study only? What was the result of the last 10 years of observation?</p> <p>4. What is the definition of "unhealthy" ageing?</p> <p>5. What is the gender score? It is mentioned in "findings to date", but it is not explained in the manuscript.</p> <p>6. Are there any relevant findings in the present follow-up? The authors explain a lot of descriptive data ("the rate of divorce had been above average"), but it is not clear, what consequence these findings have on outcome of the older patients/people.</p>
-------------------------	---

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Giulia Ferrannini, Karolinska Institute

Comments to the Author:

The authors performed the GendAge study, focusing on sex and gender differences, in a cohort of

patients from the BASE-II study.

This study had 2 phases:

- a recruitment phase (2009-2014)
- a follow-up phase (2018-2020)

including two different age groups, if my understanding is correct:

- an older age group (1671 at baseline and 1100 at follow up with complete medical information)
- a younger age group (500 or 600 at baseline - not clear - and 64 at follow up)

I found this population selection not easy to understand. Moreover, it was not possible to complete the follow up of the younger participants due to COVID-19, and accordingly, there are no results described for them nor there is a table, but some scattered information is given in the text.

I wonder whether it wouldn't be better to describe just the older population or, alternatively, to wait to have the younger population assessments completed.

We would like to thank Dr. Ferrannini for reviewing our ms and for her constructive comments. The Berlin Aging Study II (BASE-II) is an ongoing observational study. GendAge focusses on sex and gender differences in the context of cardio-metabolic diseases, and the older group of BASE-II participants was followed-up medically and psychosocially as part of GendAge. The study design of BASE-II is indeed quite complex and therefore a detailed and understandable description is what we intended to do. Following the suggestion of Dr. Ferrannini (reviewer#1) we have revised all parts of the manuscript in this sense. We also removed most of the information on the younger group (text and figure 1) which we started to assess when the pool of older participants was almost exhausted, but then data collection had to stop because of the Corona pandemic. We now mention this younger group only in a paragraph subheaded with "Other measurements" along with the note, that further details about this group will be described elsewhere (page 12).

The methods are well described and accurate, specifically as regards the gender-centered questionnaire, and there is a lot of useful information provided, which adds to the value of this study.

Thank you for this positive evaluation of our work!

However, I think that the difference between gender-related variables and sex-related assessments should be more clearly distinguished, e.g. in two different paragraphs; this would help the reader to distinguish between the two sets of variables and encourage their proper classification. There are factors, indeed, which are hard to uniquely attribute to gender or sex, but still, this should be discussed.

In the follow-up assessment part of the GendAge study, we have collected variables related to socio-cultural characteristics that are currently used to construct a gender score (GS-II). This approach will follow-up on earlier work of ours in which we had constructed a gender score in retrospect from variables that had been collected before at baseline without the intention to construct such a score. This work has already been published (Nauman et al., 2021, PMID: 33461607). We now provide some more information on the GS-I and the plans on GS-II in the manuscript, which has been restructured throughout.

Another issue is that, if I understood correctly, the authors are describing the older group only (Table 1), as in the younger group follow-up was not completed. This causes some confusion in the manuscript and does not really help to focus on the major issue.

We thank you for this comment. As described above, we have now widely removed information on the younger group that had been scattered throughout the manuscript, with the exception of a short paragraph in the "Other measurements" section (page 12).

The results should be more thoroughly described and, even though this is just a cohort profiling, might start to be put in relation to previous findings on the matter.

BMJ Open's instructions for preparing a Cohort Profile provide for the presentation of initial results in summary form and under the heading "Findings to date" (the information provided with the author guidelines: "Include a short explanation of the most notable results from the cohort so far, with references to relevant publications. This section should summarize rather than present results.")

In fact, we had partially included this summary under the heading "cohort description". Presumably, this has led to some ambiguities, which we think have been resolved with the restructuring of the entire text.

A major limitation is that the numbers are quite low. Surely, to make a great number of detailed assessments in very large populations is quite challenging, but this should nevertheless be highlighted in the limitations.

Whether the size of the study population is a limitation certainly depends on the perspective from which one looks at it. Indeed, Sohaib Haseeb (Reviewer#3) wrote: "Strengths of this paper include the broad and comprehensive nature of the data collection, large sample size constituting healthy volunteers, and comparison of age- gender- and sex-related differences."

With evaluations based on N=1,671 older subjects from the BASE-II baseline assessments, we have been able to successfully address a variety of different questions in recent years and successfully published these data (<https://www.base2.mpg.de/157731/publications>).

This will certainly also be the case with the N=1,100 data sets for whom we have two waves of within-person change data on central biomedical variables, as described in the current manuscript. There are research questions for which the study size is limiting, e.g. in the field of genetics/ large GWAS analyses where usually very small effect sizes have to be detected. However, even in this area we have been able to successfully contribute with our data in the past by participating in meta-analyses organized by large consortia. Listing N=1,100 explicitly as a limitation does therefore not seem justified to us.

Reviewer: 2

Dr. Mats Holmberg, Karolinska Universitetssjukhuset Comments to the Author:

To begin: I am not familiar with this type of manuscript, in fact, this is my first encounter with a manuscript of this type. But as I understand it, it should be a thorough description of a cohort and what the plans and considerations are for this cohort. This is what I have had in mind while reviewing.

I also want to stress that I am not a native English speaker.

Both these circumstances may colour my views.

Having said this my first suggestion will be to have this manuscript language checked. I believe many readers will be unfamiliar with the structure of the manuscript and correct English will enhance understanding.

We have revised the English language throughout the manuscript.

The second will be: stay with your object, describe the collection of the cohort, how was it done, with what purpose, how will it be investigated and what are the advantages and disadvantages of the choices you've made? I understand that you already have started to perform the analysis you planned to do but reporting results in this type of manuscript is, unfortunately, confusing. I will not go into much detail now since I feel that you need to make a choice on what type of manuscript you want to write and stick to that.

We would like to thank Dr. Mats Holmberg (reviewer#2) for this comment. BMJ Open's instructions for preparing a Cohort Profile provide for the presentation of initial results in summary form and under the heading "Findings to date". In fact, we had partially included this presentation under the heading "cohort description". Presumably, this has led to some ambiguities, which we hope to have resolved with the restructuring of the entire text.

When commenting I use the paging on the top of each page.

Specific comments

Title: By the title, it seems to me that you are presenting two cohorts? Or a study and a cohort. Or

should the title be: Cohort profile: the "older BASE II" I will come back to this issue many times.

Which cohort do you want to present and my guess is that you want to present the creation of the "older BASE II", a subsample of the BASE II study. Or the "GendAge study cohort"? the follow-up of a subsample from the BASE II study.

We understand that the title was not clear enough and changed it to: "Cohort Profile: Follow-up of a Berlin Aging Study II (BASE-II) subsample as part of the GendAge study".

Abstract

Purpose: Are you presenting a study (a protocol) or a cohort?. To me this is confusing. It seems that you start by presenting a study.

Participants: Is this the cohort? A total of 1,100 participants (older BASE-II group) This is coherent with Table 1.

Strengths and limitations

I think this section should be about the strengths and limitations of how you have designed your cohort, it includes too little of this and too much of other information now.

We absolutely agree and have rewritten most of the abstract and key points of strengths and limitations.

## Introduction

What is the reason for the new data collection?

P6r58: you are here referring to the Base II cohort but at the same time it seems as you are referring to the “older BASE II”. If you want to report on different cohorts I think it is very important that you use different names consistently.

We are thankful that the reviewer raised this point. The study design of BASE-II is indeed quite complex and therefore a detailed and understandable description is what we intended to do. We have revised all parts of the manuscript in this sense. We also removed most of the information on the younger group (text and figure 1) which we started to assess when the pool of older participants was almost exhausted. We now mention this younger group only in a paragraph subheaded with “Other measurements” along with the note, that further details about this group will be described elsewhere (page 12).

P7r3: This new instrument seems interesting and I would appreciate a more thorough description of the instrument and its advantages and disadvantages

We have now extended the paragraph on the two gender scores, the recently published retrospective version (Nauman et al., 2021, PMID: 33461607) based on baseline data, and the one, which we plan to calculate, based on the data assessed as part of the follow-up with the gender questionnaire (page 7/8 & 11).

P8r13 The numbers “the remaining pool of 1,428 subjects out of the originally 1,671” is a little confusing since they don’t match the previously described numbers.

P8r28: So the cohort was not only formed from the BASE II? You added 17 more? See below.

P8r40: the younger ones? I thought this cohort was the older ones? Confusing

We have revised the manuscript throughout to make the content more understandable in terms of the group studied and participant numbers, including revision of the flow chart (Figure 1).

P9r6: When was the follow-up? How long time? You mention how long time it took but what was the plan?

When & How long time is now mentioned under the new subheading “Contact procedure – Follow-up assessments”. We additionally added the following: “As a result of a 4-week pilot phase, we reduced the maximum number of participants examined on each of the first two study days from 6 to 4, with an interval of usually 7 days between study visit one and two. Largely because of this early adjustment, follow-up examinations lasted 21 months instead of the 15 months originally planned.” (page 7).

## Findings to date

I am sorry, but this section does seem like a study report, and as I understand this type of publication it should describe the cohort, how it was collected and what you plan to do with it. So I simply don't think this section belongs in this article.

As described above we have now moved parts of the text from the "Findings to date" section to the section describing the cohort. We now adhere to the BMJ Open's instructions for preparing a Cohort Profile by presenting initial results briefly in summary form in the "Findings to date" section.

Table 1 – I find this table clear in the description of the characteristics of the cohort. Under p-value, I think it's better just to use  $<0.05$  or  $< 0.001$  instead of publishing extremely low numbers. They are of no practical value.

We have changed this accordingly.

Table 2 – I don't think the plan for the three visits was described before?

Thank you for the important advice. We have now better introduced the study day three in the text (page 7).

Figure 1 – I find it hard to understand the relationship between the 17 and the rest in the BASE II cohort. You mention 1671+500 who were medically assessed. That is 2171. 29 are missing up to 2200. Was the 17 part of these 29? What about the remaining 12?

We can definitively understand that the small-scale presentation of the numbers can be confusing for the reader. Several institutions are involved in BASE-II, all of which collected baseline data. It is not so relevant to the reader whether medical data from the baseline survey are not available for 17 subjects in follow-up, etc., but that data from the baseline survey of at least one of the other BASE-II partner sites (e.g., genetics, psychology, socioeconomics, immunology) are also available for all N=1100 from the cohort described here. In our view, a slightly simplified and more general description of the participant numbers makes the overall text more comprehensible. In this sense and as stated above, we have revised the manuscript throughout to make the content more understandable in terms of the group studied and participant numbers, including revision of the flow chart (Figure 1).

Reviewer: 3

Mr. Sohaib Haseeb, James Cook University Comments to the Author:

The GendAge study collects follow-up data from the BASE-II study focusing on cardiovascular and metabolic diseases in a broad cohort of volunteers from Berlin, Germany. An array of data are being collected, including baseline characteristics, echocardiography data, laboratory parameters, and psychosocial and cognitive data. A gender score has also been reportedly developed.

Strengths of this paper include the broad and comprehensive nature of the data collection, large sample size constituting healthy volunteers, and comparison of age- gender- and sex-related differences. The introduction appropriately frames the study.

We would like to thank Sohaib Haseeb for the appreciative and constructive comments on our work.

Limitations of this paper include the broad range of outcomes being analyzed. It is unclear what the specific objectives are and which data have been collected and are presented as a finding. As such, it tedious to weigh the entire piece together.

We have reorganized the manuscript in order to allow a better understanding of the study goals as a whole, the details of the follow-up assessments, and the findings to date.

Specific Comments:

In general, the section “Findings to Date” does not present the data completely. Instead, it appears that more of a methodology is presented rather than findings to date. It is unclear which data are presented and which are still being collected. Specifically:

Page 9, Lines 21-34: A description of the gender questionnaire and score are given. The score and/or the elements constituting it have not been presented. Neither has it been applied and validated in the current population. Therefore, I am confused as to how this applies to the current study.

We are thankful that the reviewer raised this point. We understand that we have not been clear enough with respect to the gender score(s). We have now extended the paragraph on the two gender scores, the recently published retrospective version (Nauman et al., 2021, PMID: 33461607) using baseline data, and the one, which we plan to calculate with the specifically designed gender questionnaire using data assessed as part of the current GendAge follow-up. (page 7/8 & 11)

The gender score 2 (GS-II) will be described more comprehensively in an upcoming research article. We have not included this in the current manuscript for several reasons including the BMJ Open author guidelines stating that the cohort profile format is not intended for presenting results comprehensively: “Include a short explanation of the most notable results from the cohort so far, with references to relevant publications. This section should summarize rather than present results.” It is in this sense that we aim to succinctly describe from a bird’s eye perspective the conceptual rationale and major findings rather than provide results in detail.

Can the elements of the score be presented and applied to the participants? A separate table can be prepared to present this data.

See above.

Page 9, Lines 21-34: It is suggested that data on atrial and LV morphology, systolic and diastolic function, and vascular stiffness were obtained. Although, I can only see the data for LVEF in Table 1 and its appropriate comparison in the text; other values have not been compared.

If possible, it would be beneficial to report other echocardiography parameters and their gender-related comparisons.

Furthermore, I am curious why ICCs between an analyst and an independent supervisor are emphasized? This is an important quality-control mechanism, but perhaps not a finding.

We agree that the info on the ICC is rather a quality-control mechanism, which not necessarily needs to be mentioned in a cohort profile format. We have removed this text passage. We now present three more echocardiography parameters in table 1 and report how these did or did not differ between older men and women in our study.

Page 10, Other measurements: Where is the data concerning the psychological questionnaire and the cognitive assessments? The authors have stated that these assessments were completed, but to my eyeball review, no data have been compared — at least in the text. Are they part of Table 1?

Thank you for this comment. One advantage of this study is that very comprehensive data were collected. Because not the entire assessment battery can be presented in Table 1, we had to be selective and in addition describe in Table 2 the biopsychosocial domains of assessments obtained as part of the GendAge study. Nevertheless, we agree with Sohaib Haseeb (reviewer#3) that the area of cognitive variables was underrepresented in Table 1, so that we have now added the results of a verbal learning test in addition to the digit symbol substitution test.

Page 11, Lines 3-5: Similarly, it is stated that data concerning participant's physical activity and sleep were collected. Where are they presented? It is unclear whether these data were part of this study. For instance, much research has been performed on sleep apnea and atrial fibrillation. If such a relationship exists here would be interesting to know.

As stated above we cannot present all variables collected in Table 1. However, we agree that it is desirable to have at least some information on physical activity included. Accordingly, we have now added a variable reflecting physical inactivity to Table 1.

The research question about the relationship between sleep apnea and atrial fibrillation sounds very interesting to us. However, according to the BMJ Open author guidelines detailed results should not be reported in the format of the cohort profile, but rather be part of separate research articles. We thank R3 for prompting us to move in the suggested direction.

Page 8-11: Is there a possibility for the authors to separate the methodology from the findings. This section is quite convoluted. It starts off by presenting relevant baseline characteristics but quickly resorts to integrating methodologic elements with the findings.

Perhaps a separate section is needed for: (1) Gender questionnaire and score description w/ its elements; (2) Echocardiographic measurements; (3) Collection of genomic samples; (4) Blood sampling; and (5) Ethics clearance.

We thank the reviewer for this comment. The other reviewers have made similar comments on the structure of the manuscript. Accordingly, we have revised all parts of the manuscript in this sense, and think that this improved the clarity of the text.

I congratulate the authors on conducting a comprehensive longitudinal study of this magnitude involving participants from the community.

We again would like to thank Sohaib Haseeb for reviewing the manuscript, the constructive comments and the positive overall assessment of our work.

Reviewer: 4

Dr. Tobias Heer, Klinikum Schwabing

Comments to the Author:

The authors describe the cohort profile of the GendAge study, which is based on a reinvestigation of the Berlin Aging Study II (BASE-II). The present study focuses on major risk factors for cardiovascular and metabolic diseases and on the development of major outcomes from intermediate phenotypes in the context of biological sex and gender differences.

1. The authors should clarify if the GendAge cohort is a subgroup of the BASE-II study, which means that it is not a different population. Is the BASE-II study separately ongoing and is the GendAge population still integrated in the BASE-II study? Is GendAge the follow-up study of BASE-II? The difference of the two studies should be clarified. What does reinvestigation mean?

We thank the reviewer for this comment. The other reviewers have made similar comments demanding more clarity on the study population. We have revised the manuscript throughout to make the content more understandable in terms of the group studied and participant numbers, including revision of the flow chart (Figure 1). We now explain more clearer that the participants investigated in GendAge were identical to participants investigated in BASE-II at baseline. The term “re-investigation” was changed into “follow-up” in order to be more consistent with the wording throughout the manuscript.

2. In the abstract the authors write that the gender score was associated with a number of [...] variables and performed better in predicting differences in a subset of variables. The authors should clarify what differences they are talking about.

We are thankful that the reviewer raised this point and we have now removed the sentence concerning the retrospective gender score (GS-I) from the abstract. This was part of restructuring the manuscript as requested by the reviewers. A text passage on the GS-I is, however, included in the “Findings to date” section, and we have now added two examples, depression and life satisfaction, which have been shown to be predicted better by GS-I compared to biological sex (page 11). These results from the GendAge study have already been published (Nauman et al., 2021, PMID:

33461607).

3. Is there also a group with intervention to enhance healthy ageing or is this an observational study only? What was the result of the last 10 years of observation?

The study does not have an intervention group. The data from the follow-up examinations described in the current manuscript will be used to answer a multitude of different research questions and thus will, among other results, reveal the changes within the study population over the course of the last 10 years. According to the BMJ Open author guidelines, detailed results reporting is not suitable for the cohort profile format.

4. What is the definition of "unhealthy" ageing?

Following the prompt, we have removed the ambiguous wording (page 6).

5. What is the gender score? It is mentioned in "findings to date", but it is not explained in the manuscript.

As suggested, we now provide more information about the gender score. In particular, in the follow-up assessment part of the GendAge study, we have collected variables related to socio-cultural characteristics that are currently used to construct a gender score (GS-II). This approach will follow-up on earlier work of ours in which we had constructed a gender score in retrospect from variables that had been collected before at baseline without the intention to construct such a score. This work has already been published (Nauman et al., 2021, PMID: 33461607). We now provide some more information on the GS-I and the plans on GS-II in the manuscript (page 7/8 & 11), which has been restructured throughout.

6. Are there any relevant findings in the present follow-up? The authors explain a lot of descriptive data ("the rate of divorce had been above average"), but it is not clear, what consequence these findings have on outcome of the older patients/people.

As explained above the cohort profile format is not intended to report comprehensive results. We therefore report a brief summary of very first findings in the "Findings to date" section and thereby adhere to the BMJ Open author guidelines stating: "Include a short explanation of the most notable results from the cohort so far, with references to relevant publications. This section should summarize rather than present results."

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Haseeb, Sohaib James Cook University, College of Medicine and Dentistry
<b>REVIEW RETURNED</b>	10-Apr-2021
<b>GENERAL COMMENTS</b>	<p>The authors have put in a considerable amount of effort into addressing the reviewers' comments. As such, the manuscript has undergone major additions.</p> <p>My comments were properly addressed with sufficient additions. The structure and content of the paper has been substantially improved in my view.</p>